Depression detection in Thai language posts based on attentive network models

Vajratiya Vajrobol[1], Unmesh Shukla[1], Amit Pundir[2], Sanjeev Singh[1] and Geetika Jain Saxena[2,*]

[1]*Institute of Informatics and Communication, University of Delhi, India*

[2]*Maharaja Agrasen College, University of Delhi, India*

Abstract

Nowadays, depression is a challenging social problem that can result in desperate situations such as suicide. There exists a strong correlation between language use and the psychological characteristics of the individual at risk of depression. This study is aimed at building models that can predict depression of an individual based on the linguistic markers of their written text in Thai language. Early detection of an individual at risk of depression in the initial stages can save many lives. Social blogs are quite popular nowadays, where people elaborate on their ideas and feelings. The current study utilized Thai social blog data to create and evaluate predictive models for the early detection of individuals at risk of depressive tendencies. The methods included traditional and ensemble machine learning, neural networks, and attention-based models. This study revealed that XLM-RoBERTa, an attention network model, outperformed traditional models in terms of accuracy (79.12%), followed by Support Vector Machine (SVM) and Bi-GRU with accuracies of 78.84% and 78.56%, respectively.

Keywords

Natural Language Processing, Transformer, Depression, Deep Learning

1. Introduction

According to the American Psychiatric Association, depression is a serious medical disorder that regularly affects people's feelings, thoughts, and behavior. It may, however, be treatable. Depression is characterized by sadness or a loss of interest in former interests. It can affect your performance at work and home and lead to various mental and physical problems [1]. In addition, the National Survey on Drug Use and Health in 2020 reported that major depressive episodes affected 21 million adults in the United States who were 18 and older, or 8.4% of all adults [2]. According to a recent WHO study, depression affects 1.5 million individuals in Thailand. Females had a higher prevalence than males, with rates of depression in females and males being 2.9% and 1.7%, respectively [3]. The study in [4] discussed that inaccurate diagnoses of depressed patients might result in serious and fatal consequences. A population with little education, low socioeconomic level, and numerous barriers to accessing health services could cause depression. Similarly, the American Psychiatric Association addressed the various causes

✉ tiya101@south.du.ac.in (V. Vajrobol); unmesh.shukla@iic.ac.in (U. Shukla); amitpundir@mac.du.ac.in (A. Pundir); sanjeev@south.du.ac.in (S. Singh); gsaxena@mac.du.ac.in (G. J. Saxena)

of depression that can significantly impact an individual's life. A recent study found a causal link between the use of social media and its adverse effects, primarily depression [5].

Health experts have shifted away from traditional interactions and moved online, creating online communities for information sharing and scaling to reach more affected populations in less time [6]. These online communities offer a great opportunity of discovering individuals at risk of mental health issues and for their diagnosis and prevention. Individuals at risk of depression often use social media to express their emotions and struggles related to mental health issues. As a result, social media is an excellent resource for locating people with depressive tendencies or who are depressed. Given the volume of data, automatic, scalable computational methods could provide timely and widespread detection of depressed individuals. It would aid in preventing many fatalities in the future and help people who genuinely need it at the right moment.

Deep Learning (DL) has recently been successfully applied to mental illness detection applications [7]. It has shown significantly better performance than traditional machine learning (ML) methods used for depression detection on social media. Though DL models are effective for depression detection, their trustworthiness and robustness are still research challenges. Capturing linguistic features from timeline-based dynamic social posts may provide a crucial hint of depressive behavior over time. Blogging platforms are one type of the social media where users can express their feelings, talk about their daily lives, or vent their feelings/emotions. Several studies have reported a variety of novel techniques to identify depression in blog posts, and improved and compared the effectiveness of various methods for identifying depression-related content in Thai social blogs [8]. In this study, the dataset downloaded from the Thai social blog was preprocessed, explored and analyzed. Finally, the model was evaluated against other baseline models. To summarize, our study makes the following key contributions:

- Data exploration and analysis: utilize the PyThaiNLP library for data preprocessing and analysis to gain insight into the data.
- Comparative evaluation: developing baseline machine learning model, neural network model, and attentive network models, and fine-tuning the models for performance comparison.
- To the best of our knowledge, this is the first study that applies XLM-RoBERTa to detect depression in Thai language texts and shows the best performance
- Extensive experiments are conducted on the Thai Depression Dataset, which shows the superiority of our proposed method when compared to baseline methods.

The necessity and novelty of conducting the current study in Thai: The majority of research on the relationship between a text's linguistic properties and its author's mental health state has been primarily conducted with English texts. According to the study in Ethnologue [9], Thai is spoken by over 60 million people and is the 33rd most widely used language in the world. However, Thai is an under-researched language in the studied context, and we are aware of no other published research in Thai on the relationship between linguistic markers of a text and its writer's personality.

The work is organized into five more sections in addition to the introduction. The second section lays out the related research work on depression detection. The dataset is presented in the third section. Then, the research methodology is discussed in the fourth section. Experimental

Sentences	Class
เหตุผลต่างๆ ที่เราพยายามยึดเหนี่ยวว่าจะมีชีวิตอยู่ต่อแม่งพังพินาศ (The reasons we try to hold on to living even all have perished.)	Depression
บันทึกปีศาจ(จิตป่วย19) ความเสียใจที่บาดลึก (Devil's Diary (Sick 19) Deep Sorrow)	Depression
นั่นเป็นเหตุผลว่าครูต้องพัฒนาศักยภาพ (That's why teachers need to improve their potential.)	Non-depression
การเขียนบทกวีทำให้เขาสื่อสารกับผู้อ่านได้โดนใจกว่า อ่อนหวานกว่าและเฉียบคมกว่า (Writing poetry allows him to communicate more with the reader. Sweeter and sharper)	Non-depression

Figure 1: Example of sentences from Thai Depression Dataset

findings and analysis are presented in the fifth section. Finally, the conclusion and future works are summarized in the final section.

2. Literature Review

There are studies that have been conducting depression detection in the Thai language. The study in [10] analyzed the factors associated with and prevalence of depression among hill tribe individuals aged 30 years and over in Thailand. Few studies have been conducted on depression detection in Thai language. The study in [11] collected data from 1,105 Facebook posts and applied Support Vector Machine (SVM), Random Forest and deep learning (DL) algorithms. It was found that DL algorithms outperformed the rest with 85% accuracy in the depression class. Another relevant Thai language study [12] collected data from Thai blog posts such as Storylog, Bloggang and Blogspot with 17,116 and 16,320 posts labeled as depressed and non-depressed, respectively. The results showed that Thai-BERT achieved the highest accuracy of 77.53%, followed by Long-Short Term Memory (LSTM) Network (76.19%).

Apart from Thai language, a depression detection study conducted on Chinese microblog [13] analyzed data from the Chinese blog posting platform Weibo and auto-constructed a depression lexicon using word2vec semantic relationship graph and label propagation algorithm. It used five classification methods: Naive Bayes, LR, Random Forest, Decision Tree, and SVM, amongst which LR achieved the highest precision of 0.76.

Deep neural networks have also been applied to detect depression. For instance, the study in [7] obtained Twitter data and labeled it as control, depression, and PTSD. The result showed that CNNMax performs the best with 87.95% accuracy with optimized embedding. The study in [14] established that Convolutional Neural Network (CNN) is a better algorithm to detect depression to extract a representation of depression from audio and video. Furthermore, some studies detected depression by focusing on emotion processing, timing, and linguistic style. The study in [15] analyzed psycholinguistic features to show that Decision Tree (DT) worked better than other machine learning algorithms to detect depression.

3. Dataset

The Thai Depression Dataset [12] is a Thai language dataset that was obtained from three online sources, namely Storylog, Bloggang, and Blogspot. The dataset has 17,116 and 16,320 posts, labeled as depressed and non-depressed respectively. The training dataset contains 12,837 depressed sentences and 12,240 non-depressed sentences. The keywords / phrases - depressed, depression, depression disorder, uselessness, failure, death, overdose, suicide, cutting, and self-harm - were used as depression indicators to identify the posts of depressed class. Posts in poetic format and English were especially excluded from the dataset. Examples of posts from the dataset are shown in figure 1

4. Methodology

In Thai language, words are often grouped together without gaps, while white spaces are used to denote the beginning and end of sentences. Therefore, tokenization for Thai language works differently as compared to English. This study used the PyThaiNLP [16] library to carry out tokenization and data preprocessing for exploratory data analysis. PyThaiNLP is a text processing and linguistic analysis Python library for Thai language. Furthermore, regular English-like tokenization was performed on the dataset before generating relevant features. As shown in figure 2, post pre-processing, the TF-IDF (Term Frequency-Inverse Document Frequency) [17] features were calculated for the training and validation datasets. TF-IDF scores for each term signify its importance in the corpus. These features were fed to traditional and ensemble machine learning models. Similarly, word embeddings of variable sizes were calculated to be used as input for each deep learning model. Three traditional machine learning models, namely Multinomial Naive Bayes (MNB), Support Vector Machines (SVM), and Logistic Regression (LR), and one ensemble classifier, namely CatBoost (CB), were trained and validated on the TF-IDF scores. Three traditional deep learning networks, namely Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) network, and Bi-Gated Recurrent Unit (Bi-GRU) [18] network, and two attentive neural networks, namely Multilingual BERT (M-BERT) [19] and XLM-RoBERTa [20] were trained and validated on word embedding.

4.1. Training traditional and ensemble machine learning models

The LR classifier uses a logistic function to model the probabilities describing the binary outcomes of each trial. The study in [6] uses LR for binary text classification. The MNB classifier is one of the classic naive Bayes variants used for text classification as it implements the naive Bayes algorithm for multinomial distributed data. Unique features of SVM provide sufficient evidence to support its use for text classification [21]. CB is an improved ensemble classifier that can also be trained on text features for text classification. The current study used the aforementioned classifiers for depression detection.
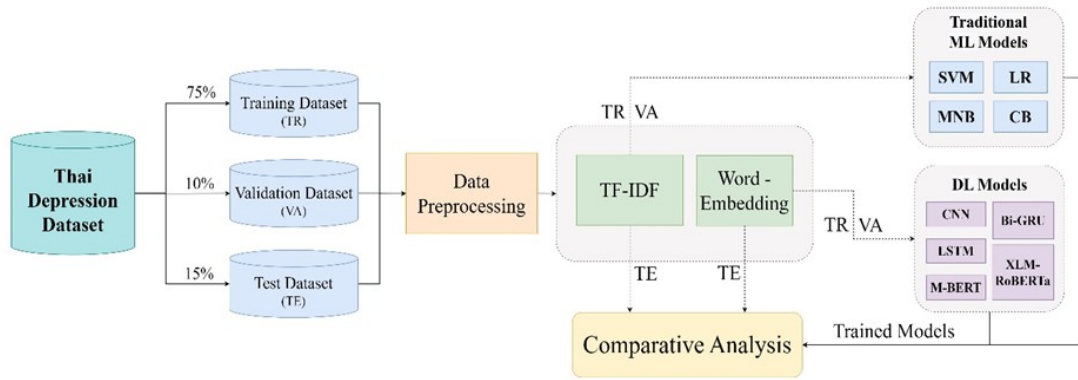
Figure 2: Framework for the development of a Thai language Depression Detector Model

4.2. Training conventional neural networks

Conventional and hybrid CNN and LSTM architectures have been used for depression detection [22]. This study trains conventional CNN, LSTM and a Bi-GRU architecture on the Thai Depression Dataset. The training of deep learning models on high-dimensional input was done using word embeddings of varying sizes [23].

4.3. Training attentive neural networks

Two transformer-based attentive neural networks - Multilingual BERT (M-BERT) and XLM-RoBERTa - were trained on the dataset for depression detection.

4.3.1. Multilingual BERT (M-BERT)

The Wikipedia entries for the top 104 languages as per the number of entries, including the Thai language, were used to train a model with the masked language modeling (MLM) objective. This pretrained BERT-based multilingual model was then trained on the Thai Depression Dataset for depression detection.

4.3.2. XLM-RoBERTa

The XLM-RoBERTa model is a multilingual version of RoBERTa [24]. It was pretrained on 2.5 TB of filtered CommonCrawl data comprising texts from 100 languages, including the Thai language. This implies that the model was built only on raw texts, with no human labeling. XLM-RoBERTa is known to be a significantly better performer when multilingual data is involved [20]. The current study used this pre-trained model for depression detection.

4.4. Evaluation Metrics

The trained models were evaluated using the common classification evaluation metrics of accuracy, precision, recall and F1-score. Furthermore, the confusion matrices were plotted and

analyzed to improve the model training. Equations (1) through (4) show the calculations for the aforementioned metrics to analyze the performance of binary classification (Depression vs Non-depression).

$$accuracy = (TP + TN)/(TP + TN + FP + FN) \tag{1}$$

$$precision = TP/(TP + FP) \tag{2}$$

$$recall = TP/(TP + FN) \tag{3}$$

$$F1 - score = 2 * (precision * recall)/(precision + recall) \tag{4}$$

5. Results and Discussion

The pre-trained XLM-RoBERTa model outperformed the rest of the models with an accuracy of 79.12%, followed by SVM and Bi-GRU, with the accuracies of 78.84% and 78.56%, respectively. The previous study [12] achieved accuracies of 75.53% and 75.06% with Thai BERT and M-BERT, respectively. As shown in Table 1, XLM-RoBERTa outperformed the rest of the models in F1-scores with a score of 0.8016. This implies that this model infers less false positives and false negatives as compared to other models. Though the precision and recall values of the transformer-based attentive neural network architectures are not as high as those of classifiers with the highest precision and recall values, the higher values of accuracy and F1-score validate the all-round performance of the attention-based classifiers. MNB and CatBoost had the highest values for precision and recall, respectively. The results in Table 1 also show that there is less variation in the performance of conventional, ensemble and neural network models trained for depression detection. CNN and LSTM were not the best in any of the performance metrics. However, among the CNN and recurrent neural network architectures, the Bi-GRU model performed the best.

The better performance of the XLM-RoBERTa model was influenced by factors such as the pretraining of the model on a number of languages and large size of the training dataset used for pretraining. This enabled the XLMRoBERTa to learn cross-language representations from a dataset comprising 100 languages. This proved helpful to learn features of the Thai language, as it is a low-resource language. Based on the capacity of the multi-language model to enhance the performance of the downstream activities, multi-language tagging data was used during the fine-tuning phase. This enabled XLM-RoBERTa to learn features of depression in Thai language texts and outperform the rest of the models.

6. Conclusion

Despite the severity of its possible repercussions, the mental condition of depression has long been overlooked. It can lead to traumatic experiences and finalities such as death by suicide. Due to the increasing influence of social media and the Internet on our personal lives, the expression of depression can be found online now more than ever. Hence, there is a dire need to

Table 1
Performance metrics of different classifiers analyzed in the current study

Algorithm	Accuracy	Precision	Recall	F1-score
MNB	0.7717	**0.8116**	0.7215	0.7639
SVM	0.7884	0.8024	0.7783	0.7902
LR	0.7823	0.7912	0.7807	0.7859
CatBoost	0.7715	0.7437	**0.8446**	0.7910
Bi-GRU	0.7856	0.7856	0.7854	7855
CNN	0.7679	0.7683	0.7673	7674
LSTM	0.7611	0.7641	0.7569	7604
M-BERT	0.7611	0.7470	0.8064	0.7756
XLM-RoBERTa	**0.7912**	0.7804	0.8239	**0.8016**

establish fine-tuned systems to detect depression in online texts, irrespective of the language in that they are written. In this study, conventional machine learning, traditional neural network, and transformer-based attentive models were used to detect depression-related content on online communities in Thai posts retrieved from Storylog, Blogspot, and Bloggang. Based on the experiments, XLM-RoBERTa, a transformer-based model, was found to be the most accurate and best-performing model with 79.12% accuracy and an F1-score of 0.8016. The other classifiers that were evaluated also performed reasonably well on all performance metrics. The current study outperformed previous studies [12] in terms of accuracy, recall, and F1-score. To the best of our knowledge, this is the first report of using XLM-RoBERTa to detect depression in Thai language. This work is limited due to less availability of text data in Thai, which in turn limits the capability of the trained models to deal with the diversity of depression expression. In the future, more data can be generated and models can be created and fine-tuned for other languages that are low-resourced. The annotation of depression on such texts can further be used to assist experts of mental health.

7. Acknowledgments

References

[1] American psychiatric association- what is depression?, 2020.

[2] National survey on drug use and health., 2020.

[3] WHO, Who south-east asia journal of public health, volume 6, issue 1, april 2017, WHO South-East Asia Journal of Public Health (2017).

[4] O. Singkhorn, T. Apidechkul, K. Pitchalard, K. Moonpanane, P. Hamtanon, R. Sunsern, Y. Leaungsomnapa, J. Thepsaw, Prevalence of and factors associated with depression in the

hill tribe population aged 40 years and older in northern thailand, International Journal of Mental Health Systems 15 (2021) 62. doi:`10.1186/s13033-021-00487-7`.

[5] M. G. Hunt, R. Marx, C. Lipson, J. Young, No more fomo: Limiting social media decreases loneliness and depression, Journal of Social and Clinical Psychology 37 (2018) 751–768. doi:`10.1521/jscp.2018.37.10.751`.

[6] A. E. Aladağ, S. Muderrisoglu, N. B. Akbas, O. Zahmacioglu, H. O. Bingol, Detecting suicidal ideation on forums: Proof-of-concept study, Journal of Medical Internet Research 20 (2018) e215. doi:`10.2196/jmir.9840`.

[7] A. H. Orabi, P. Buddhitha, M. H. Orabi, D. Inkpen, Deep learning for depression detection of twitter users, Association for Computational Linguistics, 2018, pp. 88–97. doi:`10.18653/v1/W18-0609`.

[8] A. Zafar, S. Chitnis, Survey of depression detection using social networking sites via data mining, IEEE, 2020, pp. 88–93. doi:`10.1109/Confluence47617.2020.9058189`.

[9] Ethonologue, What are the top 200 most spoken languages?., 2022.

[10] C. Chomchoei, T. Apidechkul, V. Keawdounglek, C. Wongfu, S. Khunthason, N. Kullawong, R. Tamornpark, P. Upala, F. Yeemard, Prevalence of and factors associated with depression among hill tribe individuals aged 30 years and over in thailand, Heliyon 6 (2020) e04273. doi:`10.1016/j.heliyon.2020.e04273`.

[11] K. Katchapakirin, K. Wongpatikaseree, P. Yomaboot, Y. Kaewpitakkun, Facebook social media for depression detection in the thai community, IEEE, 2018, pp. 1–6. doi:`10.1109/JCSSE.2018.8457362`.

[12] M. Hämäläinen, P. Patpong, K. Alnajjar, N. Partanen, J. Rueter, Detecting depression in thai blog posts: a dataset and a baseline, 2021.

[13] G. Li, B. Li, L. Huang, S. Hou, Automatic construction of a depression-domain lexicon based on microblogs: Text mining study, JMIR Medical Informatics 8 (2020) e17650. doi:`10.2196/17650`.

[14] L. He, M. Niu, P. Tiwari, P. Marttinen, R. Su, J. Jiang, C. Guo, H. Wang, S. Ding, Z. Wang, X. Pan, W. Dang, Deep learning for depression recognition with audiovisual cues: A review, Information Fusion 80 (2022) 56–86. doi:`10.1016/j.inffus.2021.10.012`.

[15] M. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, A. Ulhaq, Depression detection from social network data using machine learning techniques, Health Information Science and Systems 6 (2018) 8. doi:`10.1007/s13755-018-0046-0`.

[16] W. Phatthiyaphaibun, K. Chaovavanich, C. Polpanumas, A. Suriyawongkul, L. Lowphansirikul, P. Chormai, PyThaiNLP: Thai Natural Language Processing in Python, 2016. URL: http://doi.org/10.5281/zenodo.3519354. doi:`10.5281/zenodo.3519354`.

[17] J. Ramos, Using tf-idf to determine word relevance in document queries (2003).

[18] R. Skaik, D. Inkpen, Using twitter social media for depression detection in the canadian population, ACM, 2020, pp. 109–114. doi:`10.1145/3442536.3442553`.

[19] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding (2018).

[20] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, V. Stoyanov, Unsupervised cross-lingual representation learning at scale, 2019.

[21] T. Joachims, Text categorization with support vector machines: Learning with many

relevant features, 1998. doi:10.1007/BFb0026683.

[22] H. Kour, M. K. Gupta, An hybrid deep learning approach for depression prediction from user tweets using feature-rich cnn and bi-directional lstm, Multimedia Tools and Applications 81 (2022) 23649–23685. doi:10.1007/s11042-022-12648-y.

[23] H. Zhou, Research of text classification based on tf-idf and cnn-lstm, Journal of Physics: Conference Series 2171 (2022) 012021. doi:10.1088/1742-6596/2171/1/012021.

[24] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, Roberta: A robustly optimized bert pretraining approach (2019).